Color Image Segmentation Using Energy Minimization on a Quadtree Representation*

Adolfo Martínez-Usó, Filiberto Pla, and Pedro García-Sevilla

Dept. Lenguajes y Sistemas Informáticos, Jaume I Univerisity
Campus Riu Sec s/n 12071 Castellón, Spain
[auso,pla,pgarcia]@uji.es, http://www.vision.uji.es

Abstract. In this article we present the results of an unsupervised segmentation algorithm based on a multiresolution method. The algorithm uses color and edge information in an iterative minimization process of an energy function. The process has been applied to fruit images to distinguish the different areas of the fruit surface in fruit quality assessment applications. Due to the unsupervised nature of the procedure, it can adapt itself to the huge variability of colors and shapes of the regions in fruit inspection applications.

1 Introduction

Image segmentation is one of the primary steps in image analysis and visual pattern recognition. Most of image segmentation techniques are application-oriented and have been developed for specific purposes although that could be applied to a wide range of particular problems. Thus, the main motivation of the developed work has been to obtain a method able to segment images of fruits for their quality classification in visual inspection processes using a computationally efficient hierarchical representation. Particularly, the application problem that has motivated this work implies the following requirements:

1. An *unsupervised method* would be needed due to manifold variables which can arise in fruit images. Thus, any prior knowledge should be avoided for the segmentation procedure.
2. The segmentation method has to be mainly based on *color* and *edge criteria*, in order to define the segmented region boundaries as accurately as possible.

To meet the above mentioned requirements, a multiresolution Quadtree (QT) structure has been chosen to support the developed method due to its computational efficiency. The algorithm we present attachs great importance to an efficient strategy to solve the problem and the image segmentation will be conditioned and orientated by the image representation adopted, that is, by the QT and by the color and edge information as a particular and robust criterion.

* This work has been partly supported by grants DPI2001-2956-C02-02 from Spanish CICYT and IST-2001-37306 from the European Union

A. Campilho, M. Kamel (Eds.): ICIAR 2004, LNCS 3211, pp. 25–32, 2004.

As part of any segmentation process, a criterion or condition is needed to know when the final segmentation has been reached. In our algorithm, the *ideal segmentation state* is achieved through a variational method that minimizes the segmentation energy.

The main contribution of the presented work is the proposed energy function that efficiently combines color (intra-region features) with edges (borders information) using a computationally efficient hierarchical representation, a QT, to guide the minimization process. The proposed framework yields satisfactory results, particulary in fruit inspection tasks.

2 Variational Image Segmentation

The goal of variational methods for image segmentation is to develop algorithms and their mathematical analysis to minimize the segmentation energy E represented by a real value. The segmentation energy measures how smooth the regions are, the similarity between the segmented image and the original one and the similarity between the obtained edges and the discontinuities of the original image.

The Mumford-Shah model [6] has been regarded as a general model within variational segmentation methods. This model looks for a piecewise smoothed image u with a set of discontinuities, edges of the original image g.

According to Mumford-Shah's conjecture, the minimal segmentation exists but it is not unique; for each image a set of minimal segmentations exists. Therefore, with the aim of minimizing the segmentation energy we can minimize the following equation, where K is the set of discontinuities in the image domain Ω representing the *edges* of g:

$$E(u, K) = \int_{\frac{\Omega}{K}} (|\nabla u(x)|^2 + (u - g)^2)dx + length(K) \tag{1}$$

Since Mumford-Shah's work, several approaches appeared that suggested modifications to the original scheme. Recent works change equation (1) in order to improve the results. In this sense, the boundary function, which is binary in the Mumford and Shah's formulation, was changed by a continuous one which obtains a clearly defined boundary in [5]. Furthermore, in [1] the authors analyze some possible generalizations of the Mumford-Shah functional for color images. They suggest that these changes accentuate different features in edge detection and restoration.

In general, formulating variational methods have several advantages:

1. A variational approach returns explicitly a measure of the quality of the segmentation. Therefore, we are able to know how good the segmentation is.
2. Many segmentation techniques can be formulated as a variational method.
3. A variational approach can be used as a quantitative criterion in order to measure the segmentation quality.
4. Finally, a variational approach provides a way to implement non-supervised processes by looking for a minimum in the segmentation energy.

3 Energy Minimization of the Quadtree Image Representation

In this work, a function to minimize the segmentation energy is proposed. With this assumption, it is important to point out that we cannot guarantee to find a global minimum but, the experimental results obtained show that the solutions are very satisfactory.

We have developed a statistically robust functional. That is, we use an energy function that takes into account any resolution or scale change producing the same segmentation results in each case.

Let u be a smoothed image and a set of discontinuities of the original image g, let R_i be a set, with $0 < i \leq \|\Omega\|$ and $R_i \subseteq \Omega$. R_i is a family of r regions in Ω such that $\bigcup_{i=1}^{r} R_i = \Omega$ and $R_i \bigcap R_j = \emptyset$ for $i \neq j$. Let B_i represent the border of region R_i, that is, $R_i' = R_i - B_i$ is the inner part of region R_i. Finally, let γ be certain very small value that avoids dividing by zero.

Thus, $\forall x \in R_i$ let us consider the function $E_i(R_i, B_i)$:

$$E_i(R_i, B_i) = \int_{R_i} (|u(x) - m_i|)\, dx + \frac{\int_{R_i'} |\nabla g(x)|}{\int_{B_i} |\nabla g(x)| + \gamma}\, dx \tag{2}$$

In the image u, $u(x)$ represents the color of an R_i element x, and m_i is a central measure for the color value of R_i. The final segmentation energy is expressed as $E(\Omega) = \sum_i E_i(R_i, B_i) + \|\Omega\|$.

The QT structure allows us to divide an image within a complete multiresolution tree representation including neighboring information. This spatial information can be further used by a clustering strategy which groups the QT leaves using color and edge information.

Let us see the following discrete version of (2) with the same nomenclature,

$$E_i(g) = k \cdot H_i + (1 - k) \cdot G_i + \lambda \cdot length(\Omega) \tag{3}$$

that returns the segmentation energy at each region, H_i and G_i are terms as follows:

$$H_i = \frac{\sum_{R_i} (D(u(x), m_i))}{\sigma_{image}} \qquad G_i = \frac{\sum_{R_i - B_i} |\nabla g(x)|}{\sum_{B_i} |\nabla g(x)| + \gamma} \tag{4}$$

Specifically, in the QT representation, R_i is the set of leaves of the QT belonging to region i and B_i represents the boundary leaves in R_i, $0 < i \leq r$, with r being the number of regions at each iteration. The function D calculates a distance between two colors (Euclidean, Manhattan, etc.). The value $|\nabla g(x)|$ returns the gradient magnitude at the pixel x. Note that the parameter $0 < k < 1$ allows to define the weight between the color and boundary information and λ allows the method to give more or less importance to the number of regions. Finally, the value σ_{image}, which is the sum of the standard deviations of each plane in the image, contributes to normalize the first term and makes the function statistically robust. Thus, in the energy functional (3) we can distinguish three components:

1. The first one takes into account the homogeneity of each region by means of a distance from each pixel to a central measure of its region.
2. The second component promotes that the gradient magnitude will be low in the interior leaves and high in the boundary ones.
3. Finally, the third term helps the function to punish a large number of regions such as the Mumford-Shah model or many other approaches about variational image segmentation [5][1].

4 The Algorithm

4.1 Use of Color Information

Any color representation could be used. Using perceptual spaces as L*a*b* or HSI or using other representations invariant to certain features, different segmentations results will be achieved. Therefore, the process here proposed can be used on any color representation although, obviously, the results obtained will depend on it.

Paying attention to color spaces, several ones have been used with different segmentations as a result. The final regions represent the most significant features of each color space. So, in the experiments carried out the following color representations have been tested in order to verify the effect of the methodology proposed using different color spaces:

– **RGB.** This space presents a high correlation among its planes but the final result have had a pleasing performance.
– **L*a*b* and L*u*v*.** Regarding to color, these are perceptual uniform spaces where the resulting regions would be proportional to the human perceptual difference between the color of each cluster.
– **Invariant features.** In order to meet specific features within a color space, we can use some of the invariant spaces described in the literature. For instance, we use HSI to take advantage of a perceptual space as well as H plane is less influenced by non-uniform illumination. We also make use of the invariant spaces described in [7] and develop a robust feature spaces discounting shadow, illumination highlights or noise.
– **Particular spaces.** We have tested several specific spaces in order to achieve the best performance when applying the method to images of fruits. All of them are transformations from the RGB input image.
  • From the results extracted of [2], we have developed an special plane $p1 = k \cdot \log(R+G-2 \cdot B)$ which tries to find the ideal separation between the standard color of a defect and the standard color of a healthy fruit regarding oranges.
  • To have invariants for the dichromatic reflection model with white illumination a new color model $l_1 l_2 l_3$ is proposed in [4]. In our case, we only use l_2 and l_3 planes in order to make a $2D$-space which avoids the highlights and represents the chroma:

$$l_2 = \frac{(R-B)^2}{(R-B)^2 + (R-G)^2 + (G-B)^2} \tag{5}$$

$$l_3 = \frac{(G-B)^2}{(R-B)^2 + (R-G)^2 + (G-B)^2} \quad (6)$$

The functional described in the previous section has a first component H_i which represents our specific constraint to each region with the color information as a criterion. This component calculates the distance between a central measure (usually the median) and each pixel that belongs to this region. So, the smaller this result, the more homogeneous the cluster is. Note that this color term has an statistically robust behavior.

4.2 Use of Edge Information

The use of edge information in the segmentation process tries to avoid the merging of regions when the color criterion is satisfied but there exists an edge between the regions. In this sense, gradient information is used to develop a boundary map which is checked as the edge criterion. This boundary map is made from the maximum gradient value found in the R, G or B planes.

As with the color information, the functional (3) has a component G_i to try to find the boundaries as accurate as possible. This component promotes regions with no edges inside, because it promotes that the gradient magnitude will be low in the inner leaves of the region and high in the boundary ones.

4.3 The Segmentation Process

The QT structure allows us to divide an image within a complete multiresolution tree representation including neighboring information. This spatial information can be further used by a clustering strategy which joins the QT leaves using color and edge information. The multiresolution process allows us to analyze the image with a coarse-to-fine sequence described as follows:

1. We construct a hierarchical structure level by level. It is important to clarify that, talking about a QT, the meaning of *level* is a set of leaves with the same size. Thus, the first level will be the first four children leaves that descend from the whole image and so on. Therefore, while the previous named QT is created, each level is revised by the functional (3) in order to revise the clustering at that resolution. Each cluster created in any level will be taken into account in the next levels. Finally, when we finish the QT construction, the *salient regions* have been detected in a coarse way.
2. Focusing the attention on the *salient regions* (the coarse ones that have been labelled), they will be taken as the most significant groupings of the image. So, we continue the process expanding each cluster by means of a region growing method where each cluster aplies the functional (3) to its neighboring regions. This second step will take care of shaping the edges of each region by color and edge criteria.

Note that we use the functional (3) described in Sect. 3 in both of the previous steps but, whereas in the first one this application is guided by a hierarchical structure in order to develop each resolution level, in the second one the application of the functional follows a region growing strategy to achieve the final regions in a more accurate way.

Before summarizing the segmentation process, it is important to point out what are the main ideas the proposed method is based on:

1. We look for different features according to the color space used. The algorithm find groups that match regions in any color space we select, however, these groups will have properties according to the *salient features* of the color space used.
2. To guide the segmentation process, the following questions have to be solved:
   a) *the way to take to continue the segmentation process.*
   b) *how long the process have to continue.*

   The first question is solved by means of a multiresolution analysis of the image with a QT structure. Multiresolution is able to decompose the image in several resolution levels developing a coarse-to-fine process from the *salient regions* to the final shapes of each region. On the other hand, the question (b) will be determined by the functional (3) described in Sect. 3. It will be minimized in a progressive way until the functional energy stops decreasing.

The whole segmentation process is summarized in the following algorithm:

1. From RGB input image, make an edge map and create the reference color image with the selected color space.
2. Construct an oversegmented representation of the image, that is, expand the QT until every square region have all pixels with the same color. After this, create an ordered list according to region sizes.
3. The algorithm computes the functional (3) for each region and its neighbor regions in an iterative sequence that may be seen as a coarse-to-fine segmentation process.
4. If the whole image energy has decreased, reorder the list of regions by size and repeat the previous step.
5. Regroup small regions.

The previous algorithm shows the steps of the segmentation strategy used, which is basically an iterative process. Each cluster is compared to all its neighboring clusters and it will be merged when the segmentation criterion is satisfied (see Sect. 3). This clustering process stops when no other merging can be performed without increasing the energy of the segmentation. Finally, small regions are ignored and merged to their most similar neighbors giving more importance to the biggest ones.

It is important to point out that regions are arranged according to their size, giving more importance to bigger regions. This represents the spatial constraint in the merging process and facilitates merging small regions with big ones.

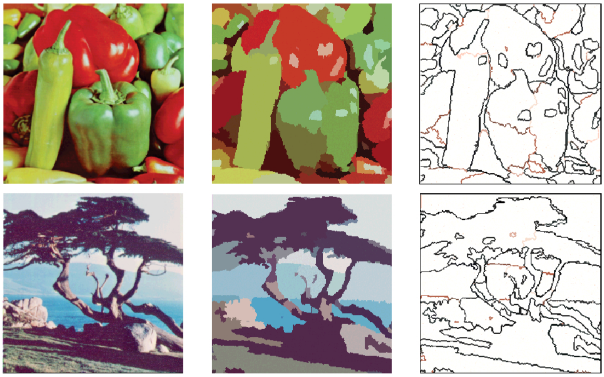

Fig. 1. Real images segmentation results.

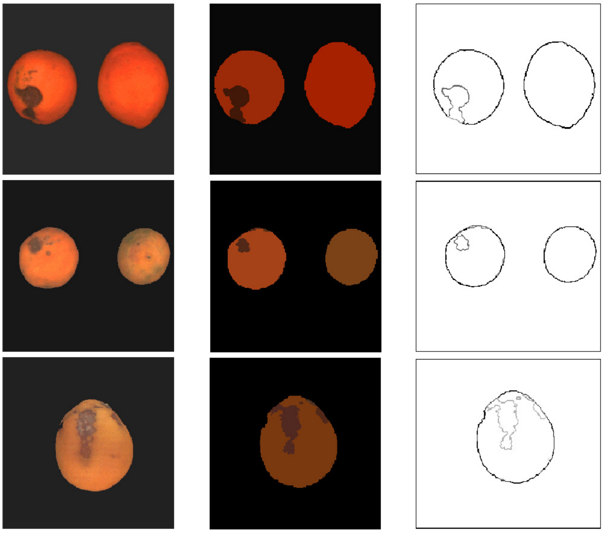

Fig. 2. Fruit images segmentation results.

5 Results

Results obtained with classical images (Fig.1) and fruit images (Fig.2) are presented. Columns represent the original images, segmented images, and the edge results, where the darker the edge line is, the greater difference between the color of neighboring regions. To show these results we have selected a perceptual color space like L*a*b* for the classical images and the transformation of equation (5) to segment the images of oranges.

Fruit image segmentation is used as input in a further process to characterize and classify the fruit surface. These visual inspection applications identify and detect different types of defects and parts of the fruit. In this sense, images in Fig. 2 show some examples of the segmentation results on different fruits and situations to be characterized. Note how the segmentation process has adapted to the regions of each image due to its unsupervised nature. For instance, the first column shows examples of fruits with various stains produced by the effect of rot and how the segmentation obtained has found the different variations of the stains of the rotten zone. This will allow the extraction of region descriptors for their classification.

Finally, it is important to point out that the algorithm has been compared with the segmentation algorithm presented in [3] which is unsupervised and employs a perceptual color space. In comparison with this algorithm, our algorithm yields similar results when tested on classical images, and outperforms it on fruit images.

6 Conclusions

In this paper, color and edge information combined with a QT representation and the minimization function has been presented. The results obtained show how the algorithm can adapt to the different situations and variability of color regions, being able to segment areas and locating the borders due to the use of gradient information during the segmentation process. Thus, this unsupervised segmentation strategy can locate color regions, and find their contours satisfactorily.

The QT representation not only guides the minimization process but also allows the segmentation at different resolution levels improving the efficiency.

References

1. A. Brook, R. Kimmel, and N.A. Sochen. Variational restoration and edge detection for color images. *Journal of Mathematical Imaging and Vision*, 18(3):247–268, 2003.
2. Y-R. Chen, K. Chao, and Moon S. Kim. Machine vision technology for agricultural applications. *Computers and Elect. in Agriculture*, (36):173–191, November 2002.
3. D. Comaniciu and P. Meer. Robust analysis of feature spaces: Color image segmentation. *IEEE Conf. Computer Vision and Pattern Recognition*, pages 750–755, 1997.
4. Theo Gevers and Arnold W.M. Smeulders. Color based object recognition. *Pattern Recognition*, (32):453–464, March 1999.
5. G.A. Hewer, C. Kenney, and B.S. Manjunath. Variational image segmentation using boundary functions. *IEEE Transactions on Image Processing*, 7(9):1269–1282, 1998.
6. D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *CPAM*, 42(4), 1989.
7. J-M Geusebroek Rein van den Boomgaard, Arnold W.M. Smeulders, and H. Geerts. Color invariance. *IEEE Transactions on PAMI*, 23(12):1338–1350, December 2001.